Module 1 Challenge Conclusion

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

* Overall, successful crowdfunding happened during the months of June and July.
* August had the most cancelled and failed crowdfunding efforts.
* Theater had the most successful crowdfunding followed by Music and then Film & Video.
* Interestingly, while theater was the most successful it also had the largest failed numbers as well.
* Plays and Rock were the most successful sub-categories.
* More than 50% of the campaigns were overall successful.
* Over a period of 10 years and in 8 different countries, people are more interested in Plays and Rock music.

2)What are some limitations of this dataset?

Limitations to this Data:

If we knew the following, we may be able to have a better understanding of the data set:

* Demographics.
* Age of people visiting the campaigns.
* Location within the country, whether the campaigns are taking place in the city or in rural areas.
* Cost of the tickets to these campaigns.

3)What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Tables and charts that would provide more clarification on the data:

* If we can add a pivot table that would have years and country in the filter. Parent category in columns and the sum of Percent funded for the campaigns that were successful, failed, cancelled and live currently that would help show us the amount of money for a certain campaign and how its success is related to the percent funded. (The table and line graph are provided in my assignment (Module-1 challenge) as Additional value 1.)
* If we can make a pivot table that would have the countries in rows and percent funded in a column. It would show us that US has a greater number of percent funded as compared to any other country in the data set. If we add the sum of campaign categories in one column. This also, tells us that US leads more campaigns as compared to any other country. (The table and bar graph are provided in my assignment (Module-1 challenge) as Additional value 2.)

Bonus Conclusion:

* Use your data to determine whether the mean or the median better summarizes the data.

The median better summarizes the data.

* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variability with successful campaigns as compared to the unsuccessful campaigns. If we refer to the table in the bonus statistical worksheet, after calculating the variance and standard deviation. It shows they are higher in number for the successful campaigns as compared to the unsuccessful. It is higher because the number of successful campaigns is more than the failed.